EXHIBIT 1

[Redacted]

Plaintiffs' [Proposed] Training Data Security and Review Protocol

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN JOSE DIVISION

In re Google Generative AI Copyright Litigation,	Master File Case No. 5:23-cv-03440-EKL-SVK Consolidated with Case No. 5:24-cv-02531-EKL-SVK
	[PROPOSED] TRAINING DATA SECURITY AND REVIEW PROTOCOL
	Judge: Hon. Eumi K. Lee Referral: Hon. Susan van Keulen, U.S.M.J.

HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY

Upon the stipulation of the parties, the following Training Data Security and Review Protocol ("Protocol") will apply to the security, disclosure, and/or review of Training Data made available by Defendants Google, LLC and Alphabet Inc. (collectively, "Google"). Plaintiffs reserve all rights to seek relief from the Court, including modification of this Protocol.

1.   For the purposes of this Protocol, "Training Data" shall be defined as data used to train relevant Google generative AI models. The Parties have and will continue to meet and confer to identify the specific Training Data sets that Google will make available to Plaintiffs. Plaintiffs sent Google a proposed list of priority datasets on May 20, 2025 (the "Priority Datasets"). Nothing in this Protocol requires any additional productions. Google reserves the right to update the Training Data made available for review should it find additional responsive data. Plaintiffs likewise reserve the right to request for inspection additional datasets. Nothing in this Protocol shall be interpreted as prohibiting the Parties from renegotiating this Protocol nor seeking leave from the Court to modify it with good cause.

2.   The "Inspecting Party" shall be defined as all Plaintiffs collectively in the above-captioned consolidated action, including their attorneys of record, agents, retained consultants, experts, and any other persons or organizations over which they have direct control.

3.   ██
████████████████████████████████████
██
██
███████████████████████████████████
██
██████████████████████████████████

4.   Training Data shall be handled and stored in accordance with the provisions of the Stipulated Protective Order entered on April 4, 2025, and any subsequent amendments or addenda thereto (the "Protective Order"). Training Data shall be designated "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" pursuant to the Protective Order, and the Inspecting Party may disclose Training Data only to those authorized to view "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" information under paragraph 7.3 of the Protective Order.

[PROPOSED] TRAINING DATA SECURITY & REVIEW PROTOCOL

HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY

1   5.   Nothing in this Protocol shall alter or change in any way the requirements of the

2 Protective Order. In the event of any conflict, however, this Protocol shall control for any Training Data

3 made available for review.

4   6.   ███████████████████████████████████████

5 ███████████████

6 ▇ ████████████████████████████████

7 ██████████████████████████████

8 ████████████████████████████████

9 ████████████████████████████████

10 ████████████████████████████████

11 ████████████████████████████████

12 ██████████████████████████████

13 ████████████████████████████████

14 ███████████████████████████████

15 ███████████████████████████████

16 ██████████████████████████████

17 ████████████████████████████████

18 ███████████████████████████████

19 ████████████████████████████

20 ████████████████████████████████

21 ███████████████████████████████

22 ████████████████████████████████

23 ████████████████████████████████

24 ▇ ████████████████████████████████

25 ██████████████████████████████

26 ████████

27 ▇ ████████████████████████████████

28 ████████

[PROPOSED] TRAINING DATA SECURITY & REVIEW PROTOCOL

HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

[PROPOSED] TRAINING DATA SECURITY & REVIEW PROTOCOL

HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY

███

██████████████████████████████████

7.   To facilitate Plaintiffs' use of the Secure Cloud Environment, Google technical personnel ("Google IT Support Personnel") shall be designated to assist Plaintiffs with setting up, operating, and addressing any technical issues with the Secure Cloud Environment. The Inspecting Party's counsel and/or experts may request that software tools and/or scripts be installed on the Secured Cloud Environment provided (a) Google approves such software tools, scripts, and/or files (and such approval not be unreasonably withheld); and (b) such other software tools, scripts, and/or files are reasonably necessary for the Inspecting Party to perform its review of the Training Data consistent with all of the protections herein. With respect to Plaintiffs' and their experts' use of the Secure Cloud Environment, the Google IT Support Personnel shall not view or access system or data logs regarding Plaintiffs' or their experts' queries or the results of queries, or the scripts or code used by Plaintiffs to run those queries. Such data logs will be automatically segregated, such that the Google IT Support Personnel shall not have access to them unless expressly authorized by Plaintiffs. In other words, neither the Google IT Support Personnel nor Google's counsel or experts shall have any visibility into Plaintiffs' work product relating to their use of the Secure Cloud Environment.[1] To ensure the security and operability of the Training Data and Secure Cloud Environment, however, the Google IT Support Personnel shall have visibility into system health and security monitoring logs to ensure the system is functioning properly, and to ensure the Secure Cloud Environment is not accessed by persons who have not been authorized to access the Secure Cloud Environment in the manner discussed above in Paragraph 6.[2]

8.   The Inspecting Party's counsel and/or experts shall be entitled to take notes relating to the Training Data, but may not copy any Training Data itself into the notes. For the avoidance of doubt, this provision shall not prevent the Inspecting Party's counsel and/or experts from recording in their notes

---

[1] Specifically, the Google IT Personnel, Google's counsel, and Google's experts shall not view or access data query information stored in the BigQuery Jobs table, *see* https://cloud.google.com/bigquery/docs/information-schema-jobs.

[2] The logging data available to the Google IT Personnel shall include Cloud audit logs (*see* https://cloud.google.com/logging/docs/audit), VPC flow logs (*see* https://cloud.google.com/vpc/docs/flow-logs), and Ops agent logs (*see* https://cloud.google.com/logging/docs/agent/ops-agent).

[PROPOSED] TRAINING DATA SECURITY & REVIEW PROTOCOL

HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY

1 particular items, files, or categories of items or files contained in the Training Data. Any notes related to

2 the Training Data will be treated as "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY."

3   9.   No copies of all or any portion of the Training Data, or other written or electronic record

4 of the Training Data, may be copied, removed, or otherwise transferred from the Secure Cloud

5 Environment except as provided herein. The Inspecting Party may obtain print outs of limited portions of

6 the Training Data to prepare court filings or pleadings or other papers (including a testifying expert's

7 expert report) by following the procedures provided herein. For purposes of this Protocol, references to

8 "print," "printing," or "print outs" are understood to refer to a Bates-stamped electronic production (as

9 described in this Paragraph). To make a request, the Inspecting Party shall create a directory entitled

10 "Takeouts" and save the desired limited portions of the Training Data in that directory. The beginning of

11 each portion of Training Data the Inspecting Party wishes to print must include the filename, file path,

12 and line numbers where the material was found in the training data or other information that allows for

13 specific identification of the material. The Inspecting Party shall alert Google when it has saved the

14 desired limited portions of the Training Data in the "Takeouts" directory that it requests to be printed.

15 Upon receiving a request, Google shall Bates number and label "HIGHLY CONFIDENTIAL –

16 ATTORNEYS' EYES ONLY" all requested pages. Within five business days from the date of request,

17 Google shall either (i) produce electronic versions to the Inspecting Party's counsel, or (ii) inform the

18 Inspecting Party that Google objects that the requested portions are excessive, not for a permitted

19 purpose, and/or not justified (see, e.g., Fed. R. Civ. P. 26(b)). In the event that Google objects, the parties

20 shall meet and confer within five business days of Google's notice of its objection. If, after meeting and

21 conferring, Google and the Inspecting Party cannot resolve the objection, the Inspecting Party shall be

22 entitled to seek a court resolution of whether the requested Training Data should be produced. In the

23 event that the Inspecting Party has designated limited portions of Training Data in the "Takeouts"

24 directory that it requests to be printed within five business days of a court-established deadline, Google

25 agrees to work with the Inspecting Party to facilitate the Bates-numbering and production set forth in this

26 Paragraph.

27   10.   All persons who will review Google's Training Data on behalf of an Inspecting Party,

28 except the Inspecting Party's counsel, must qualify under paragraph 7.3 of the Protective Order as an

[PROPOSED] TRAINING DATA SECURITY & REVIEW PROTOCOL

HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY

individual to whom "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" information may be disclosed, and must sign both the Acknowledgement and Agreement to Be Bound attached as Exhibit A to the Protective Order and the Acknowledgement and Certification attached as Exhibit A to this Protocol. All persons who review Google's Training Data in the Secure Cloud Environment on behalf of an Inspecting Party shall also be identified in writing to Google at least two business days in advance of the first time that such person reviews such Training Data.

11.   Notwithstanding any provisions of this Protocol or the Protective Order, the Parties reserve the right to amend this Protocol either by written agreement or Order of the Court upon showing of good cause.

12.   Defendants shall bear the costs associated with setting up and operating the Secure Cloud Environment, as well as the costs associated with storing and querying the Training Data in the Secure Cloud Environment, for the Priority Datasets identified by Plaintiffs on May 20, 2025. For additional datasets and models beyond the Priority Datasets identified by Plaintiffs on May 20, 2025, the Parties will meet and confer concerning costs.

IT IS SO ORDERED.

Dated: _____

_____
HON. SUSAN VAN KEULEN
UNITED STATES MAGISTRATE JUDGE

HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY

EXHIBIT A

<u>ACKNOWLEDGMENT AND CERTIFICATION</u>

I, _____ [print or type full name], of

_____ [print or type full address], declare under

penalty of perjury that I have read in its entirety and understand the Training Data Security and Review

Protocol ("Protocol") that was issued by the United States District Court for the Northern District of

California on _____ [date] in the case of *In re Google Generative AI Copyright Litigation*, Case

No. 5:23-cv-03440-EKL-SVK.

I agree to comply with and to be bound by all the terms of this Protocol, and I understand and

acknowledge that failure to so comply could expose me to sanctions and punishment in the nature of

contempt.

I further agree to submit to the jurisdiction of the United States District Court for the Northern

District of California for the purpose of enforcing the terms of this Protocol, even if such enforcement

proceedings occur after termination of this action.

I hereby appoint _____ [print or type full name] of

_____ [print or type full address and

telephone number] as my California agent for service of process in connection with this action or any

proceedings related to enforcement of this Protocol.

Date: _____

City and State where sworn and signed: _____

Printed name: _____
                        [printed name]

Signature: _____
                        [signature]

[PROPOSED] TRAINING DATA SECURITY & REVIEW PROTOCOL